

# *The Evolution of Neural Signal Decoding Techniques in Brain-Computer Interfaces from Traditional Methods to Deep Learning*

**Yiyang Zhang**

*Beijing-Dublin International College, Beijing University of Technology, Beijing, China  
yiyang.zhang@ucdconnect.ie*

**Abstract:** Brain-computer interfaces (BCIs) enable direct communication between the brain and external devices by decoding neural signals, effectively bypassing traditional peripheral pathways. Traditional BCI systems primarily relied on machine learning pipelines involving preprocessing, feature extraction, and classification. However, these approaches often faced challenges in generalizing across subjects and sessions. In recent years, the integration of deep learning has greatly enhanced neural signal decoding, enabling end-to-end modeling and improving adaptability. This paper explores the evolution of decoding techniques from traditional methods to state-of-the-art deep learning models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs) like long short-term memory networks (LSTMs), Transformer architectures, and self-supervised learning frameworks. Moreover, it discusses critical challenges such as low signal quality, high individual variability, limited data availability, and real-time processing constraints. The results suggest that deep learning methods in BCI neural signal decoding address traditional limitations by boosting accuracy and real-time performance. However, challenges like signal quality, individual variability, and real-time processing constraints persist. Thus, future developments at the convergence of artificial intelligence and BCI technologies highlight the imperative to integrate ethical and societal considerations with ongoing technological innovation.

**Keywords:** Brain-computer interface, Neural signal decoding, Deep learning, Convolutional neural networks

## **1. Introduction**

By decoding neural signals such as electroencephalography (EEG), Brain-Computer Interface (BCI) allows direct communication between the brain and external devices, bypassing traditional muscle and nerve pathways. The shift of BCI from research to practical applications in neurorehabilitation and assistive control is being accelerated by advancements in neuroscience, artificial intelligence, and sensor technology. Most current studies focus on decoding non-invasive signals such as EEG and ECoG, but issues like signal non-stationarity, low signal-to-noise ratio, and high inter-subject variability remain critical challenges limiting BCI performance [1]. Moreover, traditional machine learning methods, such as support vector machines (SVM), linear discriminant analysis (LDA), and

k-nearest neighbors (KNN), rely on handcrafted features. Although they perform well in early-stage tasks, they show clear limitations in complex tasks and cross-subject generalization, and a unified and efficient decoding paradigm is still lacking. In light of this, by adopting a literature review and systematic comparative classification method, the paper examines the development of neural signal decoding methods in BCI systems, focusing on the shift from traditional machine learning to deep learning and evaluating models' decoding performance, spatiotemporal modeling, generalizability, and effectiveness in overcoming key challenges. Specifically, it explores modeling differences of various decoding methods applied to EEG and ECoG data, and provides an in-depth analysis of deep learning architectures such as convolutional neural networks (CNN), recurrent neural networks (RNN), Transformers, and generative adversarial networks (GAN) in optimizing feature extraction and temporal modeling. Moreover, key strategies addressing cross-subject variability, low signal quality, and limited training data are summarized, as these issues critically affect neural decoding performance and limit BCI applications in complex human-computer interactions. This research promotes BCI application in complex settings and supports future technological and industrial development, offering strong academic and practical value.

## 2. Overview of neural signal decoding methods in Brain-Computer Interfaces

In BCI systems, the accurate decoding of neural signals constitutes the core process for identifying user intent and enabling human-machine interaction. For years, the decoding of neural signals has been structured around traditional machine learning approaches, which involve sequential steps of pre-processing, feature extraction, and classification [2]. Through the handcrafted feature extraction methods, raw EEG or ECoG signals are transformed into structured numerical features, which serve as inputs for classifiers tasked with inferring user intent. Traditional machine learning approaches were widely adopted in early BCI systems owing to their simplicity, computational efficiency, and interpretability. However, prior studies have pointed out that such methods often rely on features that lack sufficient scalability and generalizability. Particularly when handling highly dynamic and non-stationary EEG signals, these methods frequently suffer from notable performance degradation in cross-task or cross-subject scenarios [3].

To tackle these challenges, deep learning techniques have attracted increasing attention in recent years for neural signal decoding in BCI. Leveraging advances in GPU computing power and the availability of large-scale neural datasets, deep learning architectures have shown clear strengths in modeling and classifying high-dimensional neural signals, thereby effectively enhancing decoding accuracy and system robustness [4]. Of the many deep learning techniques available, CNNs were among the earliest to gain widespread adoption due to their effectiveness in extracting spatial structural features from multichannel EEG or ECoG signals. Classical architectures like EEGNet introduced depthwise separable convolutions, hence enabling lightweight models that maintain high performance and fit well with low-channel, resource-constrained BCI systems. Afterward, RNNs and LSTMs were applied to temporal modeling tasks in neural signals, including attention state estimation and emotion recognition. Recently, Transformer-based models have gained attention for EEG sequence modeling due to their powerful global attention mechanisms, which enable capturing long-range dependencies and improving feature representation efficiency. Additionally, emerging frameworks like GANs and self-supervised learning improve BCI by enhancing data augmentation, few-shot learning, and cross-task transfer. Thus, deep learning has expanded neural signal decoding capabilities and is becoming the dominant method in BCI research.

### 3. The neural signal decoding approach based on artificial intelligence techniques

#### 3.1. The application of CNNs in EEG decoding

In recent years, CNN has been widely adopted for EEG decoding due to its ability to automatically extract spatial features from multichannel EEG data, leveraging structural advantages like local connectivity, weight sharing, and multi-level feature extraction. This makes it particularly suitable for BCI applications such as motor imagery, attention recognition, and emotion classification [2]. Unlike traditional methods relying on handcrafted features and separate classifiers, CNN provides end-to-end modeling that directly learns high-level, discriminative features from raw signals. The convolutional layer, as the core component of CNN, functions by sliding small filters over the input data to extract local features. Each filter has learnable weights that detect specific patterns, enabling CNN to flexibly capture complex spatial structures and effectively model multichannel neural signals [5]. For instance, EEGNet introduces deep and separable convolutions to effectively extract both spatial and spectral features from EEG signals while maintaining a lightweight structure and high classification accuracy [6]. During training, CNN updates its weights through backpropagation by calculating gradients of the loss function with respect to each parameter, thereby minimizing overall prediction error. In cross-subject decoding, CNN models combined with data augmentation have achieved robust performance, with an average classification accuracy of 73% and resilience to variations in time window settings [7]. In addition, it effectively decodes EEG signals to identify motor intentions for controlling electromyographic assistive devices in upper-limb rehabilitation. With strong spatial feature extraction and good generalizability across individuals and tasks, CNN stands as one of the most representative and valuable deep learning models in EEG decoding, forming the foundation for portable, high-precision BCI systems [7].

#### 3.2. The application of RNN and LSTM in temporal decoding

In EEG decoding tasks, a considerable amount of essential information is embedded in the temporal structure of the signals. This is particularly evident in tasks such as motor imagery, attention switching, and imagined speech, where neural activity typically exhibits strong temporal dependencies. Thus, RNN and its variant LSTM have been widely employed to model the dynamic evolution of neural signals, thus improving decoding accuracy and system stability [4]. By incorporating recurrent connections, RNNs can retain information from previous time steps, enabling effective modeling of short-term dependencies. Nevertheless, standard RNNs often struggle with vanishing gradients when capturing long-term dependencies. As such, LSTMs regulate information flow through gating mechanisms. In particular, the forget gate discards irrelevant information to prevent outdated states, the input gate controls the addition of new information, and the output gate determines the relevant information for the final network output. This gating mechanism creates a “selective memory” for each EEG segment, enabling the model to retain key neural features while filtering out noise and artifacts. For example, LSTM, combined with three traditional machine learning methods, was applied to classify five vowel sounds and silence. Among these, the highest accuracy of 86.9% was achieved by LSTM, while traditional methods reached only about 60% [8]. Accordingly, RNN and LSTM architectures excel at capturing temporal dependencies in EEG signals, thus making them ideal for modeling the complex dynamics of neural activity. Their core design principles have paved the way for advanced models and remain valuable in EEG decoding.

### 3.3. The application of transformer in sequence modeling

The Transformer model, initially introduced in natural language processing, has gradually surpassed traditional sequence models because of its capability to capture long-range dependencies and its efficient parallel computing. In recent years, the model has demonstrated outstanding performance in EEG decoding tasks across multiple domains, including motor imagery, emotion recognition, and epilepsy detection [9]. By leveraging its self-attention mechanism, it effectively captures temporal dependencies and spatial features in EEG signals. In contrast to conventional sequential models, Transformers eliminate the need for fixed window sizes and bidirectional processing limitations. Instead, it establishes dependencies across the entire sequence via self-attention, which provides a significant advantage when dealing with highly nonlinear and non-stationary EEG signals. Besides, the Transformer architecture can represent multi-channel EEG data as token sequences, explicitly modeling inter-channel relationships and capturing co-activation patterns across cortical regions. Studies demonstrate that Transformers outperform CNNs and RNNs in both classification accuracy and cross-subject generalization for diverse BCI tasks, establishing them as a dominant approach in modern EEG decoding research [10]. Based on this framework, EEG Conformer integrates CNN's local feature extraction with Transformer's global modeling. It begins by capturing fine spatial patterns using one-dimensional and channel-wise convolutions, followed by multi-head attention to model long-range dependencies, and concludes with feature classification through fully connected layers. Evaluations on multiple public datasets show that it achieves state-of-the-art accuracy in motor imagery and emotion recognition tasks [11]. This is attributed to the Transformer's use of attention mechanisms, which replace fixed kernels and recurrence, enabling dynamic information selection and improved interpretability. For multi-channel EEG, the model combined with channel attention can capture latent spatial structures and connectivity, enhancing cross-regional integration [9]. Despite challenges like overfitting and large model size on small EEG datasets, Transformer performance has been improved through lightweight models, pre-training, augmentation, and transfer learning [10]. Thus, Transformer provides a novel EEG decoding paradigm, uniting CNN and RNN strengths with global context modeling, poised to lead next-gen BCI systems.

### 3.4. The application of GAN and self-supervised learning in Brain-Computer Interface

In BCI systems, training high-performance decoders often requires large amounts of labeled EEG data. However, EEG data collection and annotation are costly and slow, while individual variability and artifacts limit model generalization and application. To enhance robustness and flexibility in brain signal modeling, recent advances leverage GAN and SSL techniques. For example, GANs are effective in artifact removal, augmenting limited EEG datasets, and aligning cross-domain data in open-world EEG decoding. Models such as CycleGAN and Reversible-GAN reconstruct "clean" signals from noisy EEG or map EEG from different sources to a unified distribution, enhancing decoder stability and generalization. However, CycleGAN performs suboptimally on low-SNR EEG signals, making Reversible-GAN a more effective variant [12]. In response, the Single-Shot Reversible GAN (SSRGAN) was proposed to learn a bidirectional mapping between corrupted and clean EEG without requiring paired training data [13]. SSRGAN is specifically applied to remove ballistocardiogram artifacts commonly encountered during simultaneous EEG-fMRI acquisition. Experimental results demonstrate that SSRGAN outperforms traditional artifact removal methods like AAS and ICA by effectively suppressing interference while preserving more neural activity information, showcasing strong artifact adaptability and cross-distribution generalization [13]. In addition, some studies have introduced pseudo-tasks during unsupervised pretraining to learn

representations from unlabeled EEG data, reducing reliance on manual labels [12]. Together, GAN and SSL provide effective training strategies for low-resource BCI systems and help pave the way for more generalizable EEG decoders.

## **4. Hurdles and resolutions for neural signal decoding in Brain-Computer Interfaces**

### **4.1. Signal quality and inter-subject variability**

In BCI systems, decoding performance depends heavily on signal quality. EEG signals significantly weaken through biological layers, producing low-quality readings, and are strongly susceptible to electrical artifacts that often overlap neural frequencies, hindering stable feature extraction [4,14]. In addition, due to significant individual differences in brain structure, neural response patterns, and cognitive strategies, EEG signals generated by the same task often exhibit strong subject-specific characteristics. This variability, along with EEG's inherent non-stationarity, causes temporal and spectral changes over time and across individuals, greatly limiting decoding model generalization across subjects and sessions [15]. To enhance signal quality, strategies target both signal processing and modeling. At the signal level, artifacts are eliminated using ICA, essential frequency bands are preserved via bandpass filtering, and transient components are addressed with wavelet transforms.

Meanwhile, model-level techniques like noise-perturbed training, adaptive filtering, and adversarial training (e.g., GAN), further improve robustness against artifacts and individual variability [16]. To tackle inter-subject variability, methods emphasize subject-specific modeling for adaptation and employ transfer learning and domain adaptation to enhance cross-subject generalization. Transfer learning enables knowledge acquired from existing subjects to be transferred to new individuals, reducing the need for large training datasets. To reduce inter-subject variance, domain adaptation employs adversarial alignment and multi-source learning, thus demonstrating strong EEG decoding performance [12].

### **4.2. Real-time performance and system latency**

In real-world BCI use, rapid system response is essential for ensuring smooth user interaction and a positive experience. For neurorehabilitation and prosthetics, BCI systems must process signals within milliseconds to enable true closed-loop control between intention and action. However, the majority of EEG decoding models are trained and tested under offline conditions. High latency, poor real-time performance, and unstable responses often occur when these systems are deployed on embedded or online interactive platforms [4]. These challenges hinder the use of deep learning under tight time constraints. In particular, the main causes of latency bottlenecks include complex model architectures, large parameter sizes, and time-consuming inference processes. Models such as CNNs, BiRNNs, and Transformers typically consume substantial resources when handling entire EEG sequences [12]. Additionally, while synchronous control schemes often yield higher accuracy, their fixed time-window design causes response delays, limiting their use in continuous control tasks like brain-controlled wheelchairs and robotic arms that require immediate feedback [17]. To reduce computational cost and enhance real-time performance, real-time decoding systems often incorporate lightweight components such as shallow CNNs, compact kernels, depthwise separable convolutions, and simplified attention mechanisms. Input mechanisms have also improved, with sliding windows and rapid decision strategies enabling models to focus on key signal segments and reduce inference delays. On the hardware side, deploying trained models on high-performance edge



devices such as FPGAs and NVIDIA Jetson significantly accelerates system response, supporting real-time BCI applications [18].

### 4.3. Data scarcity and cross-task learning

Though deep learning has greatly improved neural decoding in BCI systems, data scarcity and poor generalization remain key obstacles to further progress. In particular, the strong non-stationarity and subject-specific variability of EEG signals lead to markedly different patterns across individuals performing the same task [19]. Moreover, the high cost of expert-labeled EEG data, coupled with the limited size, class imbalance, and task constraints of public datasets, poses a major challenge to effective deep model training. Collectively, these limitations undermine decoding robustness across different tasks and individuals. To enhance generalization, data augmentation techniques generate synthetic variations, like noise injection, temporal jittering, spectral shifts, and segment resampling, to increase data diversity and mitigate overfitting. Few-shot learning and meta-learning approaches incorporate structural priors and fast adaptation mechanisms, enabling models to perform well even under limited sample conditions. Besides, self-supervised and unsupervised pretraining strategies leverage auxiliary tasks, such as temporal prediction or surrogate objectives, to extract informative representations from unlabeled EEG data, thereby reducing the need for large labeled datasets in downstream tasks. Also, transfer learning and domain adaptation enhance decoder robustness across tasks and subjects by aligning cross-domain features for effective knowledge transfer [12].

## 5. Future outlook on the integration of AI and BCIs

The transition of BCI technology from laboratory research to real-world implementation is being accelerated by recent breakthroughs in artificial intelligence. By enhancing decoding accuracy and speed, CNNs and RNNs push BCI performance forward. Meanwhile, transfer and self-supervised learning ease data requirements and enhance generalization, thus paving the way for practical BCI deployment. In neurorehabilitation, these advances have driven progress in motor recovery, speech support, and epilepsy monitoring. Looking ahead, AI-powered BCIs are expected to enable precise neural modulation and personalized therapy, enhancing both rehabilitation outcomes and patient well-being. Besides, BCIs are enabling smarter environments, real-time cognitive monitoring, and more intuitive human-computer interaction. For instance, real-time EEG-based cognitive state monitoring enables adaptive driving systems to adjust assistance strategies dynamically, enhancing both safety and user experience. However, widespread adoption of BCI technologies still faces major challenges. EEG signals remain highly susceptible to noise and inter-individual variability, and current algorithms lack robustness across diverse contexts and tasks. At the same time, handling sensitive neural data raises acute privacy and ethical concerns; without clear legal frameworks, misuse could erode user rights and public trust. To bring BCIs into real-world use, AI integration must therefore deliver stronger algorithmic robustness, lighter models, stricter privacy protection, and explicit ethical safeguards. These advancements must be underpinned by collaborative research, regulatory alignment, and increased public awareness. Only such balanced growth of technology and governance will allow BCIs to become reliable tools for intelligent interaction and precision neurorehabilitation, laying the groundwork for a trustworthy human-AI symbiosis.

## 6. Conclusion

This paper examines the progression of neural signal decoding techniques in BCI systems, focusing on the transition from traditional machine learning to deep learning and the resulting changes in performance and application scope. The results show that deep learning outperforms traditional methods in accuracy, generalization, and real-time performance, establishing itself as the dominant paradigm in neural decoding. Various architectures exhibit unique strengths in capturing spatial, temporal, and long-range dependencies. Meanwhile, emerging approaches such as self-supervised learning and GANs offer promising solutions to challenges such as data scarcity and inter-subject variability. Looking ahead, the advancement of BCI depends not only on continuous algorithmic innovation but also on coordinated efforts in policy, ethics, and public acceptance. Integrating AI with BCI while ensuring privacy and data security will be critical to enabling its broader adoption and real-world impact.

## References

- [1] Li, R.L., et al. (2023). A Spectral-Ensemble Deep Random Vector Functional Link Network For Passive Brain–Computer Interface. *Expert Systems With Applications*, 215, 119341.
- [2] Saeidi, M., et al. (2021). Neural Decoding Of EEG Signals With Machine Learning: A Systematic Review. *Brain Sciences*, 11(11), 1525.
- [3] Cui, S.H., Lee, D.W., & Wen, D. (2024). Toward Brain-Inspired Foundation Model For EEG Signal Processing: Our Opinion. *Frontiers In Neuroscience*, 18, 1507654.
- [4] Craik, A., He, Y., & Contreras-Vidal, J.L. (2019). Deep Learning For Electroencephalogram (EEG) Classification Tasks: A Review. *Journal Of Neural Engineering*, 16(3), 031001.
- [5] Bradley, J., et al. (2025). Non-Invasive Brain–Computer Interfaces: State Of The Art And Trends. *IEEE Reviews In Biomedical Engineering*, 18, 26-45.
- [6] Vernon, J., et al. (2018). EEGNet: A Compact Convolutional Neural Network For EEG-Based Brain–Computer Interfaces. *Journal Of Neural Engineering*, 15(5), 056013.
- [7] Zaim, T., et al. (2025). Machine Learning- And Deep Learning-Based Myoelectric Control System For Upper Limb Rehabilitation Utilizing EEG And EMG Signals: A Systematic Review. *Bioengineering*, 12(2), 144.
- [8] Kobayashi, K., & Morooka, K. (2020). Vowel Imagery Decoding Toward Silent Speech BCI Using EEG And LSTM-Based Neural Network. *Applied Sciences*, 10(17), 5662.
- [9] Vafaei, E., & Hosseini, M. (2025). Transformers In EEG Analysis: A Review Of Architectures And Applications In Motor Imagery, Seizure, And Emotion Classification. *Sensors*, 25(5), 1293.
- [10] Pfeffer, M.A., Ling, S.S.H., & Wong, J.K.W. (2024). Exploring The Frontier: Transformer-Based Models In EEG Signal Analysis For Brain–Computer Interfaces. *Computers In Biology And Medicine*, 167, 108705.
- [11] Song, Y.H., et al. (2023). EEG Conformer: Convolutional Transformer For EEG Decoding And Visualization. *IEEE Transactions On Neural Systems And Rehabilitation Engineering*, 31, 710-719.
- [12] Chen, X., et al. (2023). Toward Open-World Electroencephalogram Decoding Via Deep Learning: A Comprehensive Survey. *ArXiv Preprint, ArXiv: 2112.06654*.
- [13] Lin, G., Zhang, S.H., & Liu, Y. (2020). Single Shot Reversible GAN For BCG Artifact Removal In Simultaneous EEG-FMRI. *ArXiv Preprint, ArXiv: 2011.01710*.
- [14] Wolpaw, J.R., & Wolpaw, E.W. (2012). Brain–Computer Interfaces: Principles And Practice. *Sensors*, 12(2), 1211-1279.
- [15] Apicella, A., Iacoviello, D., & Spezialetti, M. (2024). Improving Cross-Subject EEG Classification Via Transfer Learning: A Review. *Neurocomputing*, 570, 127438.
- [16] Fu, Z. Y., et al. (2022). Deep Learning Model Of Sleep EEG Signal By Using Bidirectional Recurrent Neural Network Encoding And Decoding. *Electronics*, 11(17), 2644.
- [17] Padfield, N., et al. (2022). A Comprehensive Review Of Endogenous EEG-Based BCIs For Dynamic Device Control. *Sensors*, 22(15), 5802.
- [18] Hosseini, M., & Khalilzadeh, M. A. (2022). A Comprehensive Survey On Deep Learning For EEG Decoding In Brain-Computer Interface Applications. *Sensors*, 22(15), 5802.

- [19] Sun, S. T., et al. (2025). Digital Therapeutics For Cognitive Impairments Associated With Schizophrenia: Our Opinion. *Frontiers In Psychiatry*, 16, 1535309.